

AD-A166 156

EXPONENTIAL BOUNDS OF MEAN ERROR FOR THE NEAREST
NEIGHBOR ESTIMATES OF EE (U) PITTSBURGH UNIV PA CENTER
FOR MULTIVARIATE ANALYSIS L C ZHAO NOV 85 TR-85-44

1/1

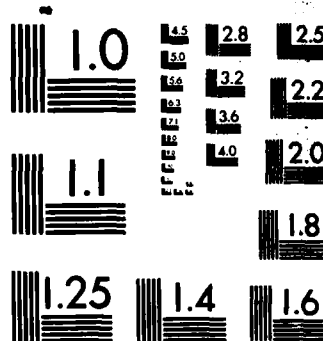
UNCLASSIFIED

AFOSR-TR-86-0065 F49620-85-C-0008

F/G 12/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFQSR-TR. 86-0065

AD-A166 156

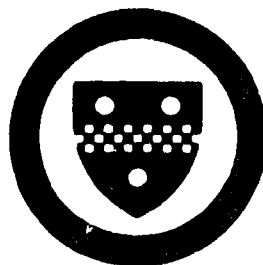
EXPONENTIAL BOUNDS OF MEAN ERROR FOR THE
NEAREST NEIGHBOR ESTIMATES OF
REGRESSION FUNCTIONS*

L. C. Zhao

DTIC FILE COPY

Center for Multivariate Analysis

University of Pittsburgh



DTIC
SELECTE
APR 1 1986
A

Approved for public release;
distribution unlimited.

2
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH - AFSC
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12.
Distribution is unlimited.
MATTHEW J. KEMPER
Chief, Technical Information Division

EXPONENTIAL BOUNDS OF MEAN ERROR FOR THE
NEAREST NEIGHBOR ESTIMATES OF
REGRESSION FUNCTIONS*

L. C. Zhao

November 1985

Technical Report 85-44

Center for Multivariate Analysis
Fifth floor, Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

DTIC
ELECTE
S APR 1 1986 D
A

*Research sponsored by the Air Force Office of Scientific Research (AFSC) under contract #F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

L. C. Zhao

Let $(X, Y), (X_1, Y_1), \dots, (X_n, Y_n)$ be i.i.d. $\mathbb{R}^r \times \mathbb{R}$ -valued random vectors with $E|Y| < \infty$, and let $m_n(x)$ be a nearest neighbor estimate of the regression function $m(x) = E(Y|X=x)$. In this paper, we establish an exponential bound of the mean deviation between $m_n(x)$ and $m(x)$ given the training sample $Z^n = (X_1, Y_1), \dots, (X_n, Y_n)$, under the conditions as weak as possible. This is a substantial improvement on Beck's result.

Key words. Regression function, nearest neighbor estimate, exponential bound, mean error, training sample.

Special

1. INTRODUCTION

Let $(X, Y), (X_1, Y_1), \dots, (X_n, Y_n)$ be i.i.d. $R^d \times R$ -valued random vectors with $E|Y| < \infty$. To estimate $m(x) = E(Y|X=x)$, the regression function of Y with respect to X , Stone (1977) and others proposed the so-called weight estimation

$$(1) \quad m_n(x) = \sum_{j=1}^n W_{nj}(x) Y_j,$$

where $W_{nj}(x) = W_{nj}(x, X_1, \dots, X_n)$ is a Borel-measurable function of its arguments.

Let V_{nj} , $j = 1, \dots, n$, be non-negative real number such that $\sum_{j=1}^n V_{nj} = 1$. For suitable-chosen metric $\|a-b\|$ on R^d (such as L_2 or L_∞), rearrange X_j , $j = 1, \dots, n$:

$$(2) \quad \|X_1^X - x\| \leq \|X_2^X - x\| \leq \dots \leq \|X_n^X - x\|$$

(ties are broken by comparing indices), and set

$$(3) \quad m_n(x) = \sum_{j=1}^n V_{nj} Y_j^X.$$

Then we obtain the nearest neighbor (NN) estimates of $m(x)$.

Many scholars studied convergence problem of these estimates from different points of view. (For the universal consistency, one can refer to, for example, Stone (1977). For the pointwise moment-consistency, see Devroye (1981). For the pointwise a.s. consistency, see Devroye (1981), Zhao and Bai (1984)). In this paper, we study another convergency of these estimates.

Write $X^n = (X_1, \dots, X_n)$, $Y^n = (Y_1, \dots, Y_n)$ and $Z^n = (X^n, Y^n)$. Let $g_n = g_n(x, Z^n)$ be an estimate of $m(x)$. In some problems, we are interested in the following mean deviation of g_n given the training sample Z^n :

$$(4) \quad \begin{aligned} D(g_n) &= E\{|g_n(X, Z^n) - m(x)| | Z^n\} \\ &= \int_{R^d} |g_n(x, Z^n) - m(x)| Q(dx), \end{aligned}$$

where Q denotes the distribution of X .

Take $k=k_n \leq n$, and put

$$(5) \quad \tilde{m}_n(x) = \frac{1}{k} \sum_{j=1}^k y_j^x.$$

For this class of estimates, Beck (1979) established the following theorem:

Suppose that the following conditions are satisfied:

- (6) (i) Y is bounded.
 (ii) $m(x)$ is continuous on R^d .
 (iii) Q has a continuous density f .
 (iv) $k \rightarrow \infty$ and $k/n \rightarrow 0$ as $n \rightarrow \infty$.

Then, for any given $\varepsilon > 0$,

$$P\{D(\tilde{m}_n) \geq \varepsilon\} \leq e^{-cn}$$

where $C > 0$ is a constant independent of n .

This theorem deals only with a special case of NN estimates, and the assumptions are rather restrictive. Recently, we substantially improved this result. We established the following:

Theorem 1. Let $m_n(x)$ be a NN estimate of $m(x)$ defined by (2) and (3).

Suppose that the following conditions are satisfied:

- (7) (i) Y is bounded.
 (ii) Q has a density f .
 (iii) There exists a sequence of integers $k = k_n$ such that

$$k \rightarrow \infty, \quad k/n \rightarrow 0, \\ \sup_n \{k \max_{1 \leq j \leq k} V_{nj}\} < \infty \text{ and } \sum_{j=k+1}^n V_{nj} \rightarrow 0.$$

Then for any given $\varepsilon > 0$, we have

$$P\{D(m_n) \geq \varepsilon\} \leq e^{-cn},$$

where $C > 0$ is a constant independent of n .

Note that the special case considered by Beck is included in this theorem. Besides, this theorem gives a substantial improvement of Beck's result, by getting rid of the continuity requirement of $m(x)$ and $f(x)$, the density of Q .

2. SOME LEMMAS.

Theorem 1 is valid for the L_2 norm or L_∞ norm on R^d , here we only give the proof for L_∞ norm. For simplicity, we make the following convention: $\varepsilon, \varepsilon_1, \varepsilon_2, \dots, C, C_0, C_1, \dots, \alpha, \beta_1, \beta_2, \delta$, etc., are all constants independent of n . I_A or $I(A)$ denotes the indicator of a set A . $\#(A)$ denotes the cardinal of set A . $S_{x,\rho} = \{u \in R^d: ||u-x|| \leq \rho\}$. Q^* and λ^* denote the outer measure generated by Q and the Lebesgue measure λ (on R^d), respectively. We need the following lemmas in the sequel.

Lemma 1 (Besicovitch Covering Lemma). Let E be bounded subset of R^d , and let K be a family of cubes covering E which contains a cube D_x with center x for each $x \in E$. Then there exist points $\{x_k\}$ in E such that

$$(i) \quad E \subset \bigcup_{x_k} D_{x_k}.$$

$$(ii) \quad \text{there exists a constant } \sigma \text{ depending only on } d \text{ such that } \sum_k I(D_{x_k}) \leq \sigma.$$

Refer to Wheeden and Zygmund (1977), pp. 185-187.

Let Q_n be the empirical measure of X_1, \dots, X_n , and $T > 0$ be a given constant. Fix $\delta \in (0, 1/2\sigma)$ and assume that $h = h_n \in (0, 1)$. Set

$$(8) \quad G_n^* = \{x \in S_{0,T}: Q_n(S_{x,h}) < \delta Q(S_{x,h})\}.$$

and

$$(9) \quad E^* = \{x \in S_{0,T}: \beta_1(2\rho)^d \leq Q(S_{x,h}) \leq \beta_2(2\rho)^d$$

for any $\rho \in (0, 1)\}$,

Where $\beta_1 > 0$ and $\beta_2 > 0$ are constants to be chosen later.

LEMMA 2. suppose that Q has a density f . Then for any $\varepsilon > 0$, we can choose β_1 small enough and β_2 large enough such that $Q^*(S_{0,T}-E^*) < \varepsilon$.

Note that for any Borel-measurable set $E \subseteq E^*$, we have

$$\beta_1 \leq f(x) \leq \beta_2, \text{ for almost all } x \in E(\lambda).$$

LEMMA 3. Suppose that Q has a density f , $h = h_n \in (0,1)$ and $nh^d \rightarrow \infty$. Then for any given $\varepsilon > 0$, we have

$$P\{Q^*(G_n^*) \geq \varepsilon\} < e^{-C_1 n}.$$

Lemmas 2 and 3 can be deduced from Lemma 1. For the proof, see Zhao (1985).

Lemma 4. Suppose that $\int_{\mathbb{R}^d} |g(x)|^p F(dx) < \infty$ for some $p > 0$, then

$$\lim_{h \rightarrow \infty} \int_{S_{x,h}} |g(u) - g(x)|^p F(du) / F(S_{x,h}) = 0$$

for almost all $x(F)$.

Refer to Wheeden and Zygmund (1977), p. 191, example 20.

3. Proof of Theorem 1

Suppose that $|Y| \leq M$. Then

$$\int \left| \sum_{j>k} V_{nj} (Y_j^x - m(x)) \right| Q(dx) \leq 2M \sum_{j>k} V_{nj} \rightarrow 0$$

as $n \rightarrow \infty$. Without loss of generality, we can assume $\sum_{j>k} V_{nj} = 0$ for any n . It is enough to prove that for each fixed $T > 0$,

$$(10) \quad P\left\{ \int_{S_{0,T/2}} |m_n(x) - m(x)| Q(dx) \geq \varepsilon \right\} < e^{-cn}.$$

By Lemma 2, there exists $\beta_i = \beta_i(\varepsilon)$, $i=1,2$, and a compact set $E \subseteq E^*$ such that

$$(11) \quad Q(S_{0,T}-E) < \epsilon/8M,$$

where E^* is defined by (9).

Fix $\delta \in (0, \frac{1}{2}\sigma)$, and take $\alpha \geq (2^d \beta_1 \delta)^{-1}$. Set

$$h = h_n = (\alpha k/n)^{1/d},$$

then $h \rightarrow 0$ and $nh^d \rightarrow \infty$ as $n \rightarrow \infty$.

By Lemma 3, there exists a compact set H_n such that with h as above

$$(12) \quad H_n \subset \{x \in S_{0,T} : Q_n(S_{x,h}) \geq \delta Q(S_{x,h})\}$$

and

$$(13) \quad P\{Q(S_{0,T}-H_n) \geq \epsilon/8M\} < e^{-C_1 n}.$$

For $x \in H_n \cap E$, $Q_n(S_{x,h}) \geq \delta Q(S_{x,h}) \geq \beta_1 \delta \lambda(S_{x,h}) = \beta_1 \delta 2^d \alpha k/n \geq k/n$, so that $x_1^x, x_2^x, \dots, x_k^x$ all fall into $S_{x,h}$.

Partition R^d into sets with the form $\prod_{j=1}^d [(i_j-1)h, i_j h)$, where $i_1, \dots, i_d = 0, \pm 1, \dots$. Call the partition Ψ . Set $\Psi' = \{B \in \Psi, B \subset S_{0,T}\}$. For $B \in \Psi'$, put

$$\tilde{W}(B) = \{B' \in \Psi, \rho(B, B') < 3h\}, \quad W(B) = \bigcup_{B' \in \tilde{W}(B)} B',$$

where $\rho(B, B') = \inf\{\|x-x'\| : x \in B, x' \in B'\}$. Then there exists a constant C_d such that for any $B \in \Psi'$ we have $\#(\tilde{W}(B)) \leq C_d$. It is easy to show by induction that, Ψ' can be divided into $C_2 (\leq C_d^2)$ disjoint subsets Ψ_i , $i=1, \dots, C_2$, such that for any two sets B_1, B_2 in the same Ψ_i , we have

$$W(B_1) \cap W(B_2) = \emptyset.$$

Denote by $B(x)$ the cube $B \in \Psi$ which contains x . If $x \in H_n \cap E$ and $B(x) \in \Psi'$, then for any $u \in B(x)$, we have $S_{x,h} \subset S_{u,2h} \subset W(B(x))$, so that, from $Q_n(S_{x,h}) \geq k/n$ it follows that x_1^u, \dots, x_k^u are also contained in $W(B(x))$. If we write

$$A_n = \{B \in \Psi^1: B \cap H_n \cap E \neq \emptyset\}$$

then, as mentioned above, for any $B \in A_n$, $W(B)$ contains the k nearest neighbors of each $x \in B$. Further, we set $H_i = A_n \cap \Psi_i$, $i=1,2,\dots,C_2$. It is easy to see that

$$\int_{S_{0,T/2}} |m_n(x) - m(x)| Q(dx) \leq \int_{S_{0,T-E}} + \int_{S_{0,T-H_n}} + \int_{H_n \cap E \cap S_{0,T/2}}.$$

By (11), we have

$$\int_{S_{0,T-E}} |m_n(x) - m(x)| Q(dx) \leq 2MQ(S_{0,T-E}) < \epsilon/4.$$

By (13),

$$\begin{aligned} P\left(\int_{S_{0,T-H_n}} |m_n(x) - m(x)| Q(dx) \geq \epsilon/4\right) \\ \leq P\{Q(S_{0,T-H_n}) \geq \epsilon/8M\} < e^{-c_1 n}. \end{aligned}$$

Hence to prove (10), it is enough to prove that

$$(14) \quad P\left\{\int_{H_n \cap E \cap S_{0,T/2}} |m_n(x) - m(x)| Q(dx) \geq \epsilon/2\right\} < e^{-c_3 n}.$$

For large n ,

$$\begin{aligned} & \int_{H_n \cap E \cap S_{0,T/2}} |m_n(x) - m(x)| Q(dx) \\ & \leq \sum_{B \in A_n} \int_{B \cap E} |m_n(x) - m(x)| Q(dx) \\ & \leq \sum_{i=1}^{C_2} \sum_{B \in H_i} \int_{B \cap E} |m_n(x) - m(x)| Q(dx). \end{aligned}$$

Put

$$\tilde{m}_n(x) = \sum_{j=1}^k v_{nj} m(X_j^x),$$

$$I_{ni} = \sum_{B \in H_i} \int_{B \cap E} |m_n(x) - \tilde{m}_n(x)| Q(dx),$$

$$(15) \quad J_{ni} = \sum_{B \in H_i} \int_{B \cap E} |\tilde{m}_n(x) - m(x)| Q(dx), \quad i=1, \dots, C_2.$$

$$\phi(B) = \int_{B \cap E} \left| \sum_{j=1}^k V_{nj} (Y_j^X - m(X_j^X)) \right| Q(dx) / Q(B \cap E),$$

$$d_{ni} = \#\{B \in H_i, \phi(B) \geq \epsilon / (8C_2)\}, \quad i=1, \dots, C_2.$$

To prove (14), it is enough to show that, for each i , $1 \leq i \leq C_2$, we have

$$(16) \quad P\{I_{ni} \geq \epsilon / (4C_2)\} < e^{-C_4 n}$$

$$(17) \quad P\{J_{ni} \geq \epsilon / (4C_2)\} < e^{-C_5 n}.$$

For almost all $x \in B \cap E(\lambda)$, $f(x) \leq \beta_2$. Hence,

$$I_{ni} \leq \epsilon / (8C_2) + 2M d_{ni} \beta_2^{\alpha k} / n.$$

Write $C_6 = \epsilon(16M C_2^{\alpha \beta_2})^{-1}$, then

$$(18) \quad P\{I_{ni} \geq \epsilon / (4C_2)\} \leq P\{d_{ni} \geq C_6 n / k\}.$$

Now we proceed to prove that, for any $B \in H_i$,

$$(19) \quad P\{\phi(B) \geq \epsilon / 8C_2 | X^n\} < e^{-C_7 k},$$

where $X^n = (X_1, \dots, X_n)$ is defined as before.

For any $\epsilon_1 > 0$ and $s > 0$, by Jensen's inequality we have

$$(20) \quad P\{\phi(B) \geq \epsilon_1 | X^n\} \leq e^{-s\epsilon_1} E\{\exp(s\phi(B)) | X^n\} \\ \leq e^{-s\epsilon_1} \int_{B \cap E} E\{\exp(s \sum_{j=1}^k V_{nj} [Y_j^X - m(X_j^X)]) | X^n\} Q(dx) / Q(B \cap E).$$

When $\{X_j^X, j \leq k\}$ is given, Y_1^X, \dots, Y_k^X are independent. From this and the inequality $|e^t - 1 - t| \leq \frac{1}{2}t^2 e^{|t|}$ for any real t , it follows that,

$$\begin{aligned} & E\{\exp(s \sum_{j=1}^k V_{nj} [Y_j^X - m(X_j^X)]) | X^n\} \\ &= \prod_{j=1}^k E\{\exp(s V_{nj} [Y_j^X - m(X_j^X)]) | X_j^X\} \\ &\leq \prod_{j=1}^k \{1 + s^2 C_8^2 k^{-2} \exp(2s C_8 k^{-1})\} \\ &\leq \exp\{s^2 C_8^2 k^{-1} \exp(2s C_8 k^{-1})\}. \end{aligned}$$

Here we have written $C_9 = \sup_n \{k \max_{j \leq k} V_{nj}\}$ and $C_8 = C_9 M$. In the same way,

$$\begin{aligned} & E\{\exp(s \sum_{j=1}^k V_{nj} [m(X_j^X) - Y_j^X]) | X^n\} \\ &\leq \exp\{s^2 C_8^2 k^{-1} \exp(2s C_8 k^{-1})\}. \end{aligned}$$

In view of (20), we get

$$P\{\phi(B) \geq \epsilon_1 | X^n\} \leq 2 \exp\{-s\epsilon_1 + s^2 C_8^2 k^{-1} \exp(2s C_8 k^{-1})\}$$

Take $s = \mu k$ with μ being small enough, we have

$$P\{\phi(B) \geq \epsilon_1 | X^n\} < e^{-C_{10} k}.$$

This is just (19).

Since for each $B \in H_i$, $W(B)$ contains the k nearest neighbors of each $x \in B$, and $W(B_1) \cap W(B_2) = \emptyset$ for any $B_1, B_2 \in H_i$, we see that when $X^n = (X_1, \dots, X_n)$ is given, $\{\phi(B), B \in H_i\}$ is a group of conditionally independent variables. Put $G(B) = \{\phi(B) \geq \epsilon_1\}$. Then by (19) and $\#(H_i) \leq \#(\Psi') \leq C_{11} n/k$, we have

$$\begin{aligned}
& P\{d_{ni} \geq C_6 n/k | X^n\} \\
& \leq P\{U_H \subset H_i, \#(H) \geq C_6 n/k \cap \bigcap_{B \in H} G(B) | X^n\} \\
& \leq \sum_{H \subset H_i, \#(H) \geq C_6 n/k} P(\bigcap_{B \in H} G(B) | X^n) \\
& = \sum_{H \subset H_i, \#(H) \geq C_6 n/k} \prod_{B \in H} P(G(B) | X^n) \\
(21) \quad & \leq \sum_{C_6 n/k \leq j \leq \#(H_i)} \binom{\#(H_i)}{j} (e^{-C_7 k})^j \\
& \leq e^{-C_6 C_7 n} \frac{\#(H_i)}{2} \leq 2^{-C_6 C_7 n} C_{11}^{n/k} \leq e^{-C_{12} n}.
\end{aligned}$$

From (18) and (21) it follows (16) is valid.

Now we proceed to prove (17). As mentioned above, for each $B \in H_i$, X_1^X, \dots, X_k^X all fall into $W(B)$. Noticing the conditions imposed on V_{nj} 's, we see that

$$\begin{aligned}
(22) \quad J_{ni} &= \sum_{B \in H_i} \int_{B \cap E} \left| \sum_{j=1}^k V_{nj} (m(X_j^X) - m(x)) \right| Q(dx) \\
&\leq C_9 k^{-1} \sum_{B \in \Psi_i} \sum_{j=1}^n I_{W(B)}(X_j) \int_{B \cap E} |m(X_j) - m(x)| Q(dx) \\
&= C_9 k^{-1} \sum_{B \in \Psi_i} \sum_{j=1}^n I_{W(B)}(X_j) Z_B(X_j),
\end{aligned}$$

where

$$(23) \quad Z_B(u) = \int_{B \cap E} |m(u) - m(x)| Q(dx) \leq 2M\beta_2 \alpha k/n.$$

Here, the following facts are used: $|m(x)| \leq M$, $f(x) \leq \beta_2$ for $x \in B \cap E$ and, $\lambda(B) \leq h^d = \alpha k/n$.

Put $\varepsilon_2 = \varepsilon(8C_2C_9)^{-1}$. To prove (17), it suffices to prove that

$$(24) \quad P\left\{ \sum_{B \in \Psi_i} \sum_{j=1}^n I_{W(B)}(X_j) Z_B(X_j) \geq 2k\varepsilon_2 \right\} < e^{-C_{13} n}.$$

Let N be a Poisson random variable with parameter n , which is independent of X_1, X_2, \dots . If $|N-n| < n\epsilon_3 = n\epsilon_2/(2M\beta_2\alpha)$, then by (23)

$$\begin{aligned} & \left| \sum_{B \in \Psi_i} \left(\sum_{j=1}^n I_{W(B)}(X_j) Z_B(X_j) - \sum_{j=1}^N I_{W(B)}(X_j) Z_B(X_j) \right) \right| \\ & \leq |N-n| 2M\beta_2\alpha k/n < \epsilon_2 k. \end{aligned}$$

It follows that

$$\begin{aligned} (25) \quad & P\left\{ \sum_{B \in \Psi_i} \sum_{j=1}^n I_{W(B)}(X_j) Z_B(X_j) \geq 2k\epsilon_2 \right\} \\ & \leq P\{|N-n| \geq n\epsilon_3\} + P\left\{ \sum_{B \in \Psi_i} \sum_{j=1}^N I_{W(B)}(X_j) Z_B(X_j) > k\epsilon_2 \right\} \end{aligned}$$

It is easy to show that

$$(26) \quad P\{|N-n| \geq n\epsilon_3\} < e^{-C_{14}n}.$$

Since $W(B)$, $B \in \Psi_i$, are disjoint, we see that for $t > 0$,

$$\begin{aligned} (27) \quad & P\left\{ \sum_{B \in \Psi_i} \sum_{j=1}^N I_{W(B)}(X_j) Z_B(X_j) > k\epsilon_2 \right\} \\ & \leq e^{-t\epsilon_2 k} \sum_{\ell=0}^{\infty} \frac{e^{-n} n^{\ell}}{\ell!} (E\{\exp(t \sum_{B \in \Psi_i} I_{W(B)}(X_1) Z_B(X_1))\})^{\ell} \\ & = e^{-t\epsilon_2 k} e^{-n} \sum_{\ell=0}^{\infty} \frac{n^{\ell}}{\ell!} \left(\sum_{B \in \Psi_i} \int_{W(B)} e^{tZ_B(u)} Q(du) + 1 - Q\left(\bigcup_{B \in \Psi_i} W(B)\right) \right)^{\ell} \\ & = \exp\{-t\epsilon_2 k + n \sum_{B \in \Psi_i} \int_{W(B)} (e^{tZ_B(u)} - 1) Q(du)\} \end{aligned}$$

Now we proceed to show that

$$(28) \quad \limsup_{n \rightarrow \infty} \sum_{B \in \Psi_i} \int_{W(B)} [\exp(\frac{n}{k} Z_B(u)) - 1] Q(du) = 0.$$

By (23), there exist constants C_{15} , C_{16} such that

$$\frac{n}{k} Z_B(u) \leq C_{15}$$

and

$$\exp\left(\frac{n}{k} Z_B(u)\right) - 1 \leq C_{16} \frac{n}{k} Z_B(u).$$

To prove (28), it suffices to show that

$$(29) \quad \limsup_{n \rightarrow \infty} \frac{n}{k} \sum_{B \in \Psi_i} \int_{B \cap E} Q(dx) \int_{W(B)} |m(u) - m(x)| Q(du) = 0.$$

Assume that $B \in \Psi_i$, $B \cap E \neq \emptyset$ and $x \in B \cap E$, then $W(B) \subset S_{x,5h}$, where $h = (\alpha k/n)^{1/d}$. By Lemma 2,

$$Q(S_{x,5h}) \leq \beta_2 (10h)^d = 10^d \beta_2 \alpha k/n.$$

Put $C_{17} = 10^d \beta_2 \alpha$, then

$$\begin{aligned} (30) \quad & \frac{n}{k} \sum_{B \in \Psi_i} \int_{B \cap E} Q(dx) \int_{W(B)} |m(u) - m(x)| Q(du) \\ & \leq C_{17} \sum_{B \in \Psi_i} \int_{B \cap E} Q(dx) \left\{ \int_{S_{x,5h}} |m(u) - m(x)| Q(du) / Q(S_{x,5h}) \right\} \\ & \leq C_{17} \int Q(dx) \left\{ \int_{S_{x,5h}} |m(u) - m(x)| Q(du) / Q(S_{x,5h}) \right\} \end{aligned}$$

By Lemma 4, for almost all $x(Q)$,

$$\lim_{n \rightarrow \infty} \int_{S_{x,5h}} |m(u) - m(x)| Q(du) / Q(S_{x,5h}) = 0.$$

Further, for $x \in S(Q)$, the support of Q , we have

$$\int_{S_{x,5h}} |m(u) - m(x)| Q(du) / Q(S_{x,5h}) \leq 2M$$

Hence, by the dominated convergence theorem, (29) is valid. Thus (28) is proved.

Take $t = n/k$ in (27), we have

$$\begin{aligned}
 (31) \quad & P\left\{\sum_{B \in \Psi_i} \sum_{j=1}^N I_{W(B)}(X_j) Z_B(X_j) > k\epsilon_2\right\} \\
 & \leq \exp\{-\epsilon_2 n + o(n)\} \leq e^{-C_{18} n}.
 \end{aligned}$$

From (25), (26) and (31), it follows that (24) holds, and (17) is valid. From (16) and (17), Theorem 1 is proved.

REFERENCES

- [1] Beck, J. (1979). The exponential rate of convergence of error for k_n -NN nonparametric regression and decision. Problems of Control and Information Theory, 8(4), 303-311.
- [2] Devroye, L. (1981). On the almost everywhere convergence of nonparametric regression function estimates. Ann. Statist., 9, 1310-1319.
- [3] Stone, C. J. (1977). Consistent nonparametric regression. Ann. Statist., 5, 595-645.
- [4] Wheeden, R. L. and Zygmund, A., Measure and Integral, Marcel Dekker, New York, 1977.
- [5] Zhao, Lincheng and Bai, Zhidong. (1984). Strong consistency of the nearest-neighbor estimates of nonparametric regression functions. Scientia Sinica (Ser. A), 27(10), 1027-1034.
- [6] Zhao, Lincheng (1985). Exponential bounds of mean error for the kernel estimates of regression functions. Technical Report (In Preparation) Center for Multivariate Analysis, University of Pittsburgh.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 86-0065	2. GOVT ACCESSION NO. ADA 166 156	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Exponential bounds of mean error for the nearest neighbor estimates of regression functions		5. TYPE OF REPORT & PERIOD COVERED Technical - November 1985
7. AUTHOR(s) L. C. Zhao		6. PERFORMING ORG. REPORT NUMBER 85-44
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis 515 Thackeray Hall University of Pittsburgh, Pittsburgh, PA 15260		8. CONTRACT OR GRANT NUMBER(s) F49620-85-C-0008
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research /NM Department of the Air Force Bolling Air Force Base, DC 20332		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 6102F 2304 A5
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE November 1985
		13. NUMBER OF PAGES 13
		16. SECURITY CLASS. (of this report) Unclassified
15. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		18a. DECLASSIFICATION/DOWNGRADING SCHEDULE
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) exponential bound, mean error, nearest neighbor estimate, regression function, training sample.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Let (X, Y) $(X_1, Y_1), \dots, (X_n, Y_n)$ be i.i.d. $R^1 \times R$-valued random vectors with $E Y < \infty$, and let $m_n(x)$ be a nearest neighbor estimate of the regression function $m(x) = E(Y X=x)$. In this paper, we establish an exponential bound of the mean deviation between $m_n(x)$ and $m(x)$ given the training sample $Z^n = (X_1, Y_1, \dots, X_n, Y_n)$, under the conditions as weak as possible. This is a substantial improvement on Beck's result.		

END

DTic

5-86